

This Survey is authorized by Title 29 USC 2

OMB 1220-0109
Expires 12/31/96

National Longitudinal Survey of Youth 1979 Cohort

Round Seventeen 1996

Questionnaire

sponsored by
U.S. Bureau of Labor Statistics

1996 NLSY79 QUESTIONNAIRE

Table of Contents

I.	Introduction	Page i
II.	Types and Formats of Question Blocks	ii
A.	R number	ii
B.	Question Name	iii
C.	Question Text	iv
D.	Distribution Code Block	iv
E.	Lead In Questions	v
F.	Default Next Question	vi
III.	Instrument Compression	vii
IV.	Save Array	viii
V.	Questionnaire	
1.	Household Interview	3
2.	Marital History, Spouse/Partner	23
3.	Regular Schooling	65
4.	Military Service	77
5.	CPS	89
6.	On Jobs	139
7.	Employer Supplement	159
8.	Periods Not Working	243
9.	Training	247
10.	Fertility	272
11.	Child Care	351
12.	Health and Job Injuries	382
13.	Income, Assets and Reciprocity	406
14.	Locator	471
15.	Interviewer Remarks	481

I. Introduction

This document describes the content of the 1996 interview of the NLSY79 cohort. This survey was conducted entirely as a computer-assisted interview. Although some cases were done over the telephone, the dominant interview mode was the in-person interview, so we will use the term CAPI (computer-assisted personal interview) to refer to the data collection technology. From 1979-1992, this cohort was surveyed using paper and pencil (PAPI) methods, and researchers could easily determine the interview content by examining the printed questionnaire and other interviewing materials.

Our goal is to make the content of the interview as comprehensible as possible by producing a clear, complete, and compact questionnaire. With this in mind, the 1996 CAPI questionnaire has been formatted as similarly as possible to a conventional questionnaire, but it also includes some additional information which enhances the usefulness of the questionnaire for data users. The resulting document accurately summarizes the “conversation” between the interviewer and respondent. Suggestions for improving the questionnaire or comments on irregularities should be directed to CHRR User Services.

We use the term “conversation” to emphasize the difference between the questionnaire and the codebook. The codebook contains the data released as a result of the interview, but some information collected during the interview is not present in the codebook. For example, the questionnaire includes questions that collect names of household members, employers, etc. Such items are not released, so the questions that collect them are not present in the codebook. Conversely, the codebook contains some information not directly collected in the interview. For example, while created variables are documented in the codebook, there is no reference to them in the questionnaire as they are not part of the conversation between interviewer and respondent.

In order to simplify the questionnaire and the public use data file, we have dropped a number of time variables from which elapsed time can be calculated (time stamps). A time stamp appears at the beginning of each major section of the questionnaire. Other time stamps have been placed around various sets of questions. Elapsed time may be very useful for a variety of methodological studies. If you are interested in these data, please contact CHRR for additional information.

To aid researchers in using the questionnaire, this introduction provides information on several topics. First, typical question blocks are described and key terms such as “R Number,” “Distribution Code Block,” and “Lead in Questions” are defined. This section explains how to follow the flow of the interview through the questionnaire. Second, the instrument compression techniques that have helped us to produce a compact questionnaire are described. Finally, this introductory section concludes with a discussion of save arrays, which provide interviewers with information specific to a given respondent (e.g., employer name, date of last interview).

II. Types and Formats of Question Blocks

Question records are the building blocks of the CAPI questionnaire, with various types of records performing different data collection and/or manipulation functions. Loosely grouped, CAPI question records will either: (a) provide administrative information that assists the interviewer or provides an introduction to the next question; (b) provide question text and control for data entry of the response; or (c) perform internal operations and calculations to guide the interview and maintain the internal data storage during the interview. All question records essential to understanding either the content of the survey or the flow of respondents through the interview will appear within the 1996 NLSY79 questionnaire.

To aid users, the 1996 questionnaire contains some elements found in the 1996 NLSY79 codebook, with codebook reference numbers provided for question items and additional information such as universe skip patterns included in each question block. Additionally, many of the elements not essential to understanding the conversational flow of the interview have been removed, making each question block less cluttered and easier to read.

Figure 1 shows the format of a typical question block; the various information fields are then explained in the following paragraphs.

R number [Question name]

Question text Question Text?
 Question text (INTERVIEWER COMMENTS)
 Question text internal code;
 Question text /* comments */
 Question text (space for continuation)

(Code text of Distribution Code Block Skip)

K1 distribution code block text1 (GO TO Rnumber)

K2 distribution code block text2

K3 distribution code block text3 (GO TO *Qk-nn)

.

.

.

Kn distribution code block textn

Lead In(s):

Default Next Question:

Figure 1. CAPI Question Block

A. R Number

The NLSY79 identifies variables by reference numbers (R numbers). This identifier for a particular variable in a particular year does not change, even as new waves are released. The R number has the format Rxxxxx.xx. Every question in the questionnaire which appears in the codebook has its reference number printed next to it as an identification. This allows the user to identify immediately which items he or she will be able to extract from the CD and to find a given question on the CD with a minimum of effort. Any item that is not in the public use data set will not have an R number in this questionnaire; these items are identified by question name only.

Most questions generate a single variable in the codebook. The R numbers for these questions will not include the decimal extension; they will simply be assigned a number with the format Rxxxxx.

Some questions generate more than one variable. For example, some questions collect information about the date an event happened and split this information into three variables: month, day and year. Questions that ask the respondent which of several responses are appropriate, with the respondent being able to pick all responses that apply, can generate a great many variables. In these questions, each possible outcome is a distinct variable, coded as selected, not selected, or a valid skip. When a question record generates multiple variables, those variables will have decimals in the R number. For example, for a date question the month might be stored as R45000.00 and the year as R45000.01. If a question asked the respondent to list all months in 1994 when he or she was in school, the question would produce 12 variables, one for each month. The first variable (January) would be R47000.00, the second variable (February) would be R47000.01, and so forth.

Examples:

R41000. This is the unique R number for this variable. This R number matches the R number in the codebook.

R42000.03 The fourth variable generated by a single question record. The first variable generated by that question record would be R42000.00, the second would be R42000.01, and so on.

B. Question Name

Question names are the unique identifiers assigned to identical questions across CAPI survey years. They replace the deck-and-column numbering system used in previous NLSY79 surveys.

Question names provide information on: (a) the section of the instrument from which each question was derived—e.g., question names beginning with Q2- come from the “Marital Status” section of the questionnaire, while Q8- questions come from the “Training” section; (b) the order in which the question was administered within each section; and (c) whether a question was part of a “looped” or repeated series. Previous and subsequent NLSY79 CAPI surveys that field this exact question will use the same question name to help identify comparable questions across years.

Example: Q6-16.3. This question comes from the “On Jobs” section of the questionnaire, which is usually the sixth section (as modules are added or dropped the number of sections in an interview changes). The “16” only suggests its original position; sometimes question 14a is inserted between 14 and 15, and sometimes questions are dropped or reordered. The “.3” indicates this question is in the third set of a repeated group of questions. In this case, this question is in a set of questions asking about the third employer the respondent reported on at the previous interview but for whom the respondent was not working at the time of the previous interview.

Identification of the multiple Employer Supplements, each of which collects information on a single employer, has been simplified. Employer Supplement questions start with “QES “ and a number identifying the employer to which a given question applies. For example, question names with the prefix QES1 indicate questions asked about the first employer, while the QES5 questions were administered concerning the fifth employer. Decimals appended to the employer number and question number denote looped or repeated questions relating to periods not working for a specific employer and pension plans from a specific employer under which the respondent may be covered. Only the first Employer Supplement is printed in this questionnaire; see the “Instrument Compression” section below for details.

If the same question was asked in 1993 and/or 1994, and again in 1996, it will have the same question name. This makes it easier for the user to link identical items across these rounds.

C. Question Text

This field contains the text of the question that the interviewer asked the respondent or other information used to direct the flow of the interview. The following types of text may appear:

Question Text:	Text that is asked of the respondent appears in a conventional mixture of upper and lowercase letters.
INTERVIEWER COMMENTS:	Text in all uppercase is either an instruction to the interviewer or a clarifying comment. NORC interviewers do not read these items to the respondent.
Internal Code:	When a question record contains a machine instruction, that instruction is printed in this field. Such instructions end with a semi-colon. These instructions are accompanied by internal comments as described below.
/* comments */	Internal comments that explain what a machine instruction does are set off by a leading “/*” and a trailing “*/.” These comments explain the function being performed by the internal code so that persons unfamiliar with the command syntax can follow the interview protocol.
Continuation:	Up to five lines of question text are allowed in each question block. When a question requires more space, it is preceded by one or more “Dummy Questions” that simply allow the question text to be as long as necessary.

D. Distribution Code Block

When a question requires the choice of one or more items from a predefined list, that list, or distribution code block, is shown in the question block. Each item in the distribution code block has three components:

Code:	The numeric code associated with each possible response is listed first. All data in the NLSY79 public use data file are numeric, so each possible item that could be selected is associated with an integer.
Text:	Next is the item text or description of the code, such as Yes or No. If the question uses a “show card” (this will be indicated in the question text field), the items on the show card are sometimes identified by letters or numbers; these letters or numbers are included with the response text in this questionnaire.
Skip:	When the selection of a specific response determines which question will be asked next, a <i>(GO TO Rxxxxx)</i> or <i>(GO TO *QK-nn)</i> notation follows the response text. If the skip instruction contains the question name preceded by an asterisk, then the next question does not appear in the public use data. That item will still appear in the questionnaire.

For some questions that collect numeric data (hours, weeks, dollar amounts, etc.), the question block contains what appears to be a distribution code block. This is present whenever the next question to be asked depends upon the numeric value falling in some range. In such cases the distribution code block is constructed so it corresponds to these skip conditions. However, the interviewer would actually enter the exact number provided by the respondent and this exact number would be recorded in the data set.

E. Lead in Questions

When using PAPI questionnaires, researchers needed to study the instrument in detail to determine under what circumstances a particular question was asked. The CAPI questionnaire simplifies this process somewhat by including a “Lead in(s)” field in the question block. This field lists the questions that lead into each particular item.

Examples:

R41000:[default]	This means that the default path from question R41000 leads to the current question, but there may be conditions under which a different path would be taken.
R41000:[all]	This means that all cases in which R41000 is asked will lead to the current question. Please note this does not imply all respondents are asked question R41000.
R41000:[1:4]	This means that whenever the response category for question R41000 takes on the values one to four (inclusive), the next question will be the current question record. If the response to R41000 is some other value, the respondent may next be asked a different question.
R41000.	When the lead in question is shown without a qualifier such as [Default], [All], or [1.4], this is a sign that the “skip pattern compression” algorithm dropped intervening records that could be suppressed (see “Instrument Compression” below). In such cases, if the user refers back to the lead in question, this earlier question will clarify the flow of the interview.
*Q11-63: (All)	This means that all cases in which Q11-63 is asked will branch to the current question. Because it has no R number, the indication is that Q11-63 is not included in the public use data set.

By tracing the skip pattern backward, one may determine the universe of respondents asked a given question. This universe information is a new feature not present in the documentation for paper and pencil interviews.

F. Default Next Question

This field specifies the question to be asked next unless another skip specification in the distribution code block applies. If the next question is in the public use data set, its R number is provided. If the next question is one which is not available to the public, the question name preceded by an “*” is shown.

III. Instrument Compression

In order to provide a more compact questionnaire, two types of records have been dropped. The following do not appear within this questionnaire: (1) “looped” questions or repetitive series; and (2) question records that perform internal operations not necessary for understanding the conversational flow of the survey.

First, repetitions of questions that are asked multiple times are not included. For example, the CAPI instrument asks questions about the education, employment, etc., of up to twenty persons in the household. Instead of printing all twenty sets of questions, only questions for the first person are printed. A note is appended reminding the user that 19 more sets of these questions are available in the questionnaire. Users can readily identify such repeated “loop” questions because the question names end with “.1”; the replications not printed in the questionnaire (which do appear in the codebook) have names that end “.2” to indicate the second replication, etc. Employer Supplements, the set of questions asked about each of the respondent’s employer arrangements, are also question loops. However, the naming convention differs slightly from that discussed above, taking the form QES1-xx, QES2-xx, and so forth. The CAPI questionnaire allows up to ten sets of Employer Supplement questions, although only supplements one through five are included in the public use data, following long-standing practice.

Questions that have been dropped because they are replications will, occasionally, appear in lists of lead-in questions or branching instructions. However, if a referenced question has been dropped, the user can easily determine its function by looking at the corresponding question in the first loop, that is, the question whose name is identical except that it has the suffix “.1.”

Second, the instrument was compressed by eliminating question records that contained only nonessential skip pattern information. The 1996 CAPI instrument included over 17,000 numbered question records, only some of which are essential to understanding the action content and conversational flow of the survey. The process of dropping unessential records and then changing the lead in and next question specifications so that the skip pattern does not have any gaps is referred to as “skip pattern compression.” The resulting questionnaire still contains all relevant information about skip patterns and universes and is easier for researchers to use.

IV. Save Array

“Save arrays” are fields in which information is stored. Each such field, or save array, is assigned a name. For example, the name of the respondent’s current employer would be stored in the “EMPLOYER.NAME(1)” save array for use during the interview. This “save array name” is then used to reference and access the data in the save array field during the actual operation of the survey. Each time a question refers to the current employer, the computer automatically inserts the name of the employer into the question. Information stored in save arrays is used:

- a) as part of the actual text of survey questions (e.g. the name of a given household member, child, employer, types of training programs, different sources of income for household members);
- b) to govern paths through the questionnaire. The save array field is accessed, and the content (or absence of content) is checked. The path through the questionnaire is determined based upon the information in the save array field;
- c) as elements in equations and calculations which are used to produce new pieces of information which may be used later in the instrument.

When used in a question text, save arrays are similar to the text fills used in the PAPI NLS interviews. These text fills indicated that the interviewer should substitute a given word or phrase into the question at the time of the interview so it read correctly. For example, in a printed PAPI questionnaire one would often see “*Since (DATE OF LAST INTERVIEW) have you...*”; this phrase would instruct the interviewer to insert the actual date for a given respondent. However, correctly inserting such text fills was the job of the interviewer and often involved flipping between pages.

One immediate difference in the CAPI questionnaire is that there are far more text fills, and most of them are automated or handled by the computer. Some CAPI text fills are identical to those used in the PAPI interviews, such as “*he/she*” to handle gender fills. The major new feature is the automation of the text fills through the save arrays. In the date of last interview example, the computer code in the CAPI questionnaire would automatically insert that date into the question. The interviewer never actually sees the text “*[Date of last interview]*” on the screen, only the date stored in that location.

In this questionnaire, text fills which refer to a save array are represented by bracketed text describing what piece of information is substituted in the interview. Although this bracketed text is never seen by an interviewer, researchers can use this information to follow the flow of the interview.